Do Cross-Section Dependence and Parameter Heterogeneity Matter? Evidence on Human Capital and Productivity in Greece

Nikos Benos and Stelios Karagiannis

University of Ioannina, Center of Planning and Economic Research

September 2013

Do Cross-Section Dependence and Parameter Heterogeneity Matter? Evidence on Human Capital and Productivity in Greece

Nikos Benos[a,b] and Stelios Karagiannis[c]

Abstract

This paper estimates production functions for Greek regions over 1981-2003, using a novel human capital dataset. We construct rich human capital series, where data for employees are decomposed according to their education level. Our empirics include recent non-stationary panel techniques, allowing for cross-section dependence and parameter heterogeneity in production technology, along with fixed effects and dynamic panel estimators. We show that ignoring cross-section dependence and parameter heterogeneity has a serious distorting impact on the estimated results, both qualitatively and quantitatively. Our evidence shows that tertiary education has a strong positive association with labor productivity growth, while secondary education exhibits a negative relationship. Primary education, public capital and net agglomeration do not display any relation with growth. Overall, findings suggest that policy makers should account for spillovers alongside technology heterogeneity and direct their efforts towards the expansion of tertiary education in poor regions in order to promote convergence.

J.E.L. Classification: J24, O47, I21, R10, C23

Keywords: Labor productivity, Growth, Human capital, Regional data, Panel data

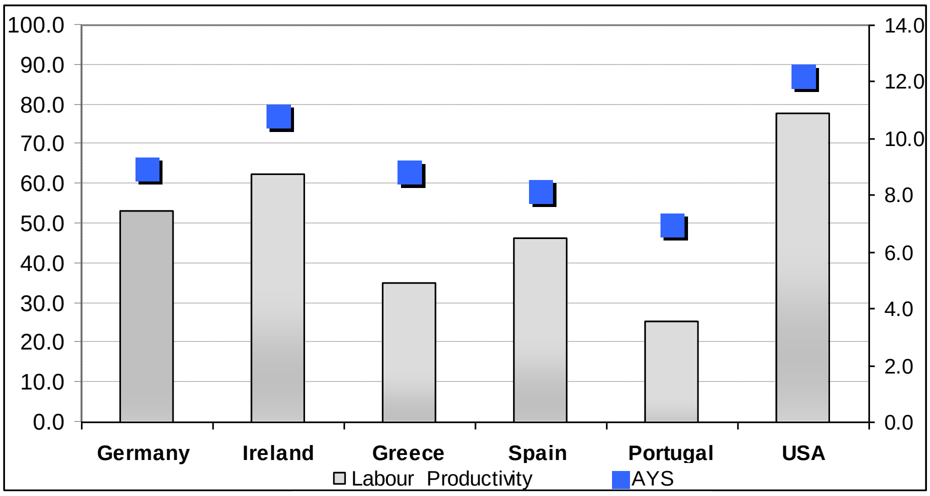
[a] Department of Economics, University of Ioannina, University Campus, P.O Box 1186, 45110 Ioannina, Greece. Email: nbenos@cc.uoi.gr
[b] Laboratory of Applied Economics and Social Policy (LAESP), University of Ioannina, University Campus, P.O Box 1186, 45110 Ioannina, Greece.
[c] Center of Planning and Economic Research (KEPE), 11 Amerikis Str., 10672 Athens, Greece. Email: stkarag@kepe.gr

1. Introduction

Recently, there has been an increasing interest in the estimation of production functions. In this framework, numerous studies have investigated labor productivity differences across countries and regions. This line of research has been motivated by persistent disparities both across and within countries worldwide and the quest for convergence in living standards across the globe set, among others, by international organizations like the United Nations (United Nations, 2000).

Following this strand of inquiry, we investigate regional disparities in terms of labor productivity in Greece, putting special emphasis on human capital. Economic theory suggests a positive relation between human capital and productivity, implying that the former constitutes a basic force behind income convergence or divergence. Theoretical contributions focus on the distinct roles of human capital accumulation, human capital stock or both mechanisms behind the growth process (Lucas, 1988; Romer, 1990, Azariadis and Drazen, 1990). However, by looking at the data for a number of developed countries, we observe that there is no apparent relation between labor productivity and human capital, measured by average years of schooling, for the 1980-2010 period (see Fig. 1 below).[1]

FIGURE 1: LABOR PRODUCTIVITY INDEX AND AVERAGE YEARS OF SCHOOLING (average 1980-2010)



Source: Labor productivity (GDP per employee; 2005=100) from European Commission's AMECO Database; Average years of schooling from Barro and Lee (2010) (right-hand axis)

[1] Data are obtained from the OECD and the Barro and Lee (2010) Educational Attainment Dataset, respectively.

There have been several attempts to test this relationship formally, usually employing cross-section country data. These studies use formal education indicators as proxies for human capital, because investment in education plays a central role in human capital accumulation. However, they provide contrasting results: growth effects of human capital are estimated to be positive, statistically insignificant or even negative in some cases (Pritchett, 2001; Barro and Sala-i-Martin, 2004). A basic reason for these puzzling results is that most studies use international datasets, but incorrectly impose equal returns to schooling (homogeneous coefficients) among countries (Temple, 1999a, 1999b; Krueger and Lindahl, 2001; Di Liberto, 2007). This is problematic, because education provision is affected by educational institutions, which often differ across countries. Moreover, returns to education are likely higher in countries with a better educated labor force (Azariadis and Drazen, 1990). Another issue is that education investment is not linked with productivity in some cases, i.e. education is not only an investment but also a consumption good for individuals. Finally, especially in less developed countries, public sector employs almost all skilled labor force, creating distortions in the estimation of education returns, since these are determined mostly by government regulations and not market forces (Griliches, 1997).

In this paper, we try to shed some light into this puzzle, by estimating the production functions for Greek regions[2] in a unified framework, putting emphasis on human capital effects emerging from different education levels. This is where our contribution lies: our study represents a novel attempt to investigate labor productivity in Greek regions for a fairly long period (1981-2003) using annual data. Here, we should note that we are the first to construct human capital data, which include graduates of primary, secondary and tertiary education as a percentage of the employed population, at NUTS 3 level. We focus on a homogeneous dataset, since Greek regions are characterized by common institutions and a harmonized education system in terms of regulatory framework. In addition, the decomposition of education into levels allows us to estimate their differential effects on productivity. Such estimations are frequently ignored by the literature. Another methodological contribution of this paper consists in the joint examination of the role of cross-section dependence and parameter heterogeneity in a non stationary panel framework. These issues arise due to common shocks, spillovers and production

[2] Increasing evidence suggests that regional rather than national economies are the decisive units at which growth takes place (Ohmae, 1995; Storper, 1997; Cheshire and Malecki, 2004).

technology heterogeneity (Costantini and Destefanis, 2009; Eberhardt and Teal, 2011). This is crucial given the high degree of interdependency among regional units. Finally, we allow for agglomeration effects on labor productivity, due to location (McDonald and McMillen, 2007). Specifically, we account for labor market pooling due to easy access of both employers and employees to alternatives and population proximity, which facilitate skill "matches" and product distribution, respectively (Cohen and Morrison Paul, 2009). Concurrently, our analysis accommodates congestion diseconomies.

We first, show that parameter heterogeneity and cross-section dependence seriously affect the empirical results, both qualitatively and quantitatively. Second, regional productivity growth exhibits a strong positive relationship with tertiary education and a weaker negative relationship with secondary education. On the contrary, we do not verify an association between labor productivity growth and primary education, as well as public investment. The evidence on secondary education is probably explained by the low skills of Greek graduates compared to their OECD peers, as shown by student performance in international standardized tests. Third, we uncover that the positive spillover effects of density are equally important to its negative congestion effects, implying the absence of net agglomeration economies. This is in contrast with most existing literature, which concludes in favor of positive net agglomeration economies.

The paper is organized as follows. Section 2 gives a review of the theoretical and empirical literature on human capital and growth. Section 3 provides the theoretical framework of our empirical model. Section 4 describes the data and the econometric methodology. In Section 5, we discuss the empirical results and Section 6 offers some concluding comments. The Appendix contains detailed information on variable definitions and data construction.

2. Literature review

The large theoretical literature on human capital and economic growth, can be summarized as follows: (i) human capital accumulation boosts growth (Lucas, 1988); (ii) growth depends on existing human capital stock, which generates new knowledge (Romer, 1990) and facilitates the imitation or adoption of foreign technologies (Nelson and Phelps, 1966); (iii) the impact of human capital depends on human capital stock accumulated within a given period (Azariadis and Drazen, 1990).

As discussed in the Introduction, the empirical literature provides mixed results as far as the effect of human capital on economic growth is concerned. One of the earliest attempts to introduce human capital in the empirical growth literature, is made by Mankiw et al. (1992), who estimate a positive output elasticity with respect to the working-age population with secondary education in 121 countries during 1960–1985. Studies employing country-level data were followed by research using regional data, similarly to our study. For instance, Arbia et al (2010) conclude that tertiary education attainment boosts growth in 271 NUTS 2 EU regions in 1991-2004 accounting for spatial effects due to institutions and geography. Soukiazis and Antunes (2011) show that secondary education attainment contributes to growth directly and indirectly through interaction with exports, in Portuguese NUTS 3 regions during 1996–2005. Abel and Gabe (2011) uncover a strong positive relationship between working-age population with a college degree and GDP per capita in 290 US metropolitan areas during 2001-2005. In addition, Pablo-Romero and Gómez-Calero (2013), using a translog production function, conclude that private physical and human capital are complementary and exhibit decreasing returns, for 50 Spanish provinces during 1985–2006.

Some recent research uses panel cointegration techniques to deal with non stationary series. For instance, Kosfeld and Lauridsen (2004) conclude that employed people with at least secondary education increase both GDP per employed and GDP per capita in 180 German labor markets in the year 2000 accounting for spatial effects. Similarly, Bronzini and Piselli (2009) estimate a positive long-run relationship between average employee schooling years and both labor productivity and output in 1985-2001 for 19 Italian regions. Karnik and Lalvani (2012) conclude that the gross enrollment ratio exhibits a strong positive effect on GDP per capita in 19 Indian states during 1981–2005 and the contribution of education to growth is larger than that of physical capital.

A few papers have examined the possibility of differential growth effects between schooling levels. For example, Asteriou and Agiomirgianakis (2001) and Sari and Soytas (2006) find a cointegrating relationship between enrollments in primary, secondary and tertiary education and GDP in Greece (1960-1994) and Turkey (1937-1996), respectively. Petrakis and Stamatakis (2002) show that primary and secondary education matter for growth in less developed countries, while tertiary education becomes important in developed economies, using data for 24 countries. According to Papageorgiou (2003), primary education is important for final goods production, while post-primary education is necessary for technology adoption and innovation in 80 countries during 1960-1987. Pereira and Aubyn (2009) find that increasing education of the working-age population at all levels except tertiary education has a positive effect on GDP per worker growth in Portugal over 1960. Ramos et al. (2010) conclude that tertiary and secondary education increase labor productivity and growth, respectively, while primary education does not exert any influence in the Spanish NUTS 3 regions during 1980–2007. Interestingly, they find negative geographical spillovers from tertiary education. Ding and Knight (2011) find that higher education enrollments have a stronger positive growth impact than secondary enrollments, while primary school enrolment has no effect in 30 Chinese provinces for 1978-2007. Finally, Cuaresma et al. (2012) use Bayesian model averaging with 48 growth determinants for 255 NUTS 2 EU regions during 1995-2005 allowing for spatial spillovers. They find that workers with higher education have a robust positive association with GDP per capita growth, while they also include in the estimations workers with secondary, primary education as well as a lifelong learning variable.

3. The model

We study the role of human capital in the growth process, adopting a production function approach. Thus, we specify our theoretical model by augmenting a standard aggregate Cobb–Douglas production function as follows:

$$Y_{it} = A_{it} K_{Git}^{a} H_{Pit}^{\beta} H_{Sit}^{\gamma} H_{Tit}^{\delta} L_{it}^{1-a-\beta-\gamma-\delta} \tag{1}$$

where, Y denotes real output of region i (i = 1, …, 51) during period t (t = 1981, …, 2003), K_{Git} the public physical capital stock, $H_{Pit}, H_{Sit}, H_{Tit}$ stand for human capital stock produced through primary, secondary, tertiary education, respectively, L_{it} is employment, and A_i is a Hicks-neutral TFP indicator.[3] Thus, we allow for differential impact of the three stages of education. The introduction of three types of human capital enables us to obtain more accurate estimates of the model's parameters.

[3] Specification (1) assumes constant returns to scale with respect to public capital, human capital and labor. We have actually tested this assumption and found that it is true for our dataset (results available upon request from the authors).

Equation (1) in "per employed worker" terms takes the form:

$$\frac{Y_{it}}{L_{it}} = A_{it}\left(\frac{K_{Git}}{L_{it}}\right)^{a}\left(\frac{H_{Pit}}{L_{it}}\right)^{\beta}\left(\frac{H_{Sit}}{L_{it}}\right)^{\gamma}\left(\frac{H_{Tit}}{L_{it}}\right)^{\delta} \tag{2}$$

Taking logs of Equation (2), we obtain Equation (3)

$$\ln\left(\frac{Y_{it}}{L_{it}}\right) = \ln A_{it} + a\ln\left(\frac{K_{Git}}{L_{it}}\right) + \beta\ln\left(\frac{H_{Eit}}{L_{it}}\right) + \gamma\ln\left(\frac{H_{Sit}}{L_{it}}\right) + \delta\ln\left(\frac{H_{Tit}}{L_{it}}\right) \tag{3}$$

We extend (3) in line with the literature on spatial agglomeration and productivity (Ciccone and Hall, 1996; Ciccone, 2002; Abel and Gabe, 2011, Abel et al., 2012). Specifically, empirical research has shown that thick labor markets imply significant productivity benefits by improving the quality of matches between workers and jobs (Andersson, Burgess, and Lane, 2007). Also, firms locating in each others' proximity may incur higher productivity. At the same time, there is the possibility for congestion diseconomies. In light of these, we assume that density (*D*) operates through the technology parameter (*A*) as follows:

$$A_{it} = B_{it} D_{it}^{\zeta} \tag{4}$$

where ζ represents the elasticity of output with respect to density and B_{it} denotes other determinants of technology which are independent of density. The parameter ζ measures the net agglomeration impact of density, which includes both the (positive) agglomeration and (negative) congestion effects due to density. Thus, the sign of ζ will depend on the relative strength of these opposing forces. Equation (3) becomes

$$\ln\left(\frac{Y_{it}}{L_{it}}\right) = \ln B_{it} + \zeta\ln D_{it} + a\ln\left(\frac{K_{Git}}{L_{it}}\right) + \beta\ln\left(\frac{H_{Pit}}{L_{it}}\right) + \gamma\ln\left(\frac{H_{Sit}}{L_{it}}\right) + \delta\ln\left(\frac{H_{Tit}}{L_{it}}\right) \tag{5}$$

Equation (5) constitutes the base of our empirical analysis.

4. Data and econometric methodology

4.1 Data

In order to investigate the relationship between education and labor productivity in a regional production function framework we employ a dataset, which includes a balanced panel of 51 regions (NUTS 3 level) over the 1981-2003 period.[4] This is the longest possible data set with consistent series currently available at this level of regional disaggregation. Our data include the number of graduates of the three levels of education, namely primary, secondary and tertiary as a percentage of the employed population per region. These data treat human capital as a stock variable. We constructed them for the first time for Greek NUTS 3 regions for such a long period. To do so, we used census information on education attainment, employment and retirement rates for the three education levels, combining it with enrollment, employment and retirement rates for all years in order to estimate education attainment for primary, secondary and tertiary education in the inter-census years, following the Barro and Lee (2010) methodology (see Appendix for details on the construction of the variables).

Furthermore, regional public capital stock is included in our regressions in order to disentangle the effect of physical capital from the effect of human capital on growth (Krueger and Lindahl, 2001), because regional private capital data at NUTS 3 level are not available for Greece. We believe this is a good proxy for capital stock at regional level, because public and private capital behave similarly in Greek regions (Louri, 1989). Note that the private-public investment correlation is 0.97 at national level.[5] We construct public capital stock using the investment perpetual method, with a depreciation rate of 5% according to the standard practice in the literature (Rovolis and Spence, 2002).[6] Specifically, our proxy is based on the actual investment included in state budget expenditures at regional level.

Regarding density, we employ three alternative measures, namely employment density, labor force density and population concentration. These correspond to different but not necessarily contrasting possibilities for density effects. Employment density equals the number of employed over the area of the respective region, labor force density uses the same formula, the only difference being that density effects take place within the whole labor force, not just employed persons. Population concentration measures density effects within the whole population; however, it measures them in relative terms, since it

[4] Table 1, available in the Appendix, provides detailed definitions of the variables used in our estimations.
[5] This estimate uses time series data (1981-2003) obtained from the Annual Macro-Economic (AMECO) database of the European Commission's Directorate General for Economic and Financial Affairs.
[6] See Table 1 in the Appendix, for the definition and a description of the construction of the variable. Please note that this is the second attempt for the computation of public capital in the relevant literature; the initial one was by Labrinidis et al. (2005).

calculates density of each region relative to the average population density of all regions (Roos, 2005). We proceed this way, because there are alternative variables suggested in the literature for testing the effects of density on labor productivity and we do not know a priori which one is relevant for Greece (Brülhart and Mathys, 2008; Ciccone, 2002; Piras et al., 2012; Abel et al, 2012; Cuaresma et al, 2012).

According to our descriptive statistics (see Table 2 in the Appendix), Greek regions are characterized by numerous disparities. Differences in regional real GDP per worker reveal large spatial labor productivity differentials; it ranges from €15,294 to €51,904€ in 1981 and €22,305 to €84,798 in 2003. Also, the shares of employees, who are graduates of the three education levels, differ substantially across regions as well as through time. Specifically, the share of workers with primary education declined substantially from 50.8% in 1981 to 31% in 2003, while the fraction of workers with secondary and tertiary education increased significantly. The former reached 43% from 16.6% and the latter 16% up from 6.7% of the employed population during the period examined. All measures of density differ notably between regions. Employment density ranges from 3 to 432.3 and labor force density from 3.7 to 472.7. Population concentration varies from 39.4 to 41.8, indicating that the most important population movements between regions had already taken place at the start of the period studied in our work. Finally, the regional allocation of real public capital per worker in Greece presents large disparities both across space and time, ranging from €8,781 to €1,112,453 in 1981 and €6,593 to €1,646,179 in 2003.

4.2 Cross-section dependence and parameter heterogeneity

To overcome the issue of spurious regression, which characterized earlier studies on the relation between education and regional productivity – due to the neglect of the time series properties– we follow a four-step approach: first, we assess the cross-section dependence of the series; second, we test them for stationarity; third, depending on the order of integration of the series, we decide whether to test for cointegration; if we do so and cointegration exists, an error correction model (ECM) is estimated, which permits to analyze the long-run relationship between the variables jointly with the short-term adjustment towards the long-run equilibrium; if we do not apply cointegration testing, we investigate the short-run dynamics of the education– labor productivity relationship.

Specifically, time-varying heterogeneity due to unobserved common shocks, which affect all units (in our case regions), introduces cross-section correlation or dependence in the error terms, which can lead to inconsistency and incorrect inference in standard panel econometric approaches (Phillips and Sul, 2003; Pesaran, 2006). At the same time, the assumption of cross-section independence is strong for regional data; cross-region co-movements of economic variables are most likely due to common shocks and spillover effects (Economides, 1996). So, we test for cross-section dependence and model it using unobserved common factors, but not spatial effects.[7] We do this, because if we use the latter method, we may not account for endogenous time-varying variables (i.e. trade, FDI and policy), which can not be simply approximated by distance measures (Pesaran, 2004). We model cross-section dependence by postulating:

$$B_{it} = e^{\eta_i + \theta_i' f_t + e_{it}} \tag{6}$$

$$x_{mit} = \pi_{mi} + p_{mi}' g_{mt} + q_{1mi} f_{1mt} + ... + q_{nmi} f_{nmt} + v_{mit} \tag{7}$$

$$f_t = \rho' f_{t-1} + e_t \text{ and } g_t = \kappa' g_{t-1} + \upsilon_t \tag{8}$$

where η_i idiosyncratic regional technology term, e_{it} idiosyncratic random shock, x_{it} right-hand side observed variables in (5), $m=1,...k$, f_t, g_t unobserved common factors and $f_{mt} \subset f_t$. If we substitute (6) in (5), denoting natural logs by lower-case letters, we get:

$$y_{it} = \beta_i' x_{it} + u_{it} = \beta_i' x_{it} + \eta_i + \theta_i' f_t + \varepsilon_{it} \tag{9}$$

The variables y_{it}, x_{it} are possibly nonstationary, β_i are region-specific elasticities of labor productivity with respect to the various inputs, θ_i, p_i, q_t are region-specific factor loadings, η_i, π_{mi} are region-specific fixed effects and $e_{it}, v_{it}, \varepsilon_{it}, \upsilon_{it}$ stand for i.i.d errors.[8] The factors f_t, g_t can be nonlinear and nonstationary. The presence of f_t in (6) - (8) induces endogeneity, because the regressors are correlated with the unobservables of the production function (u_{it}). If we do not account for f_t, g_t during estimation, we will produce biased estimates of β_i and incorrect inferences (Pesaran, 2006). Additionally, if these factors are nonstationary, estimation approaches neglecting heterogeneous common

[7] Baltagi (2008) in a detailed review of the panel unit root and cointegration literature, points towards the vital importance of controlling for cross-sectional dependence.

[8] Here, we note that the above common factor specification can not discriminate among possible channels of cross-section dependence.

factors do not identify β_i (Kapetanios et al., 2011; Eberhardt and Bond, 2009).

The above empirical production function framework allows for parameter heterogeneity across regions in the impact of observables (inputs) and unobservables (TFP) on output. "New growth theory", which emphasizes production functions differences across cross-sections, justifies heterogeneous technology parameters (Durlauf et al., 2001). It argues that production technology heterogeneity may mean that countries can choose an 'appropriate' production technology from many possible options.

Also, our empirical framework deals appropriately with any business cycle effects, i.e. idiosyncrasies of regional economies or global shocks with heterogeneous impacts (Chudik et al., 2011). Our model is intended for use with annual data, enabling us to deal properly with their time-series and cross-section properties (Eberhardt and Teal, 2011).
To test for cross-section dependence, we apply the CD test by Pesaran (2004), which uses the correlation coefficients between the time series for each panel member. In our case, for *N=51* regions, this would be the 51 x 50 correlations between region *i* and all other regions, for *i=1* to *N-1*. Denoting the estimated correlation between the time-series for region *i* and *j* as $\hat{\rho}_{ij}$, the Pesaran CD statistic is given by:

$$\sqrt{\frac{2T}{N(N-1)}}\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\hat{\rho}_{ij}\right) \tag{10}$$

where T is the time-series dimension of the panel. Under the null of cross-section independence, the above statistic follows the standard normal distribution. The statistic is robust to nonstationarity, parameter heterogeneity and structural breaks and performs well even in small samples. Results are reported in Table 1 below. The test statistics imply decisive rejection of the cross-section independence hypothesis for all variables. Therefore, it provides strong evidence that cross-section dependence exists for them. Given that the observed variables are correlated across regions, it is natural to expect that the unobservables, which are contained in the error term, will also be correlated across regions.

[INSERT TABLE 1 HERE]

4.3 Unit root tests

In the long run, series such as output or capital stock often display strong persistence, so it is reasonable to test if our series are 'non-stationary' processes (Nelson and Plosser, 1982; Lee *et al.*, 1997; Pedroni, 2007). This is important, because in the case of a non-stationary variable the addition of observations does not help to learn about its distribution, e.g. its mean, variance etc., since they do not converge to constant values. As time-averaging does not alter the order of integration, any macro production function is likely to include at least some cross-sections with non-stationary input and output variables, and their time series properties must be taken into account to avoid bias and/or inefficiency (Granger, 1988; Granger and Siklos, 1995).

So, we examine if our data are stationary. If the variables are not stationary in levels, we check for stationarity in differences. The recent literature has concluded that inference on the time series properties of the data can be improved when applying integration and cointegration tests to the whole panel rather than to each unit separately. This way we address the problem of low power of conventional time series tests, since we increase the sample size considerably.

We apply the Pesaran (2007) panel unit root test for multiple variables and lags in models with and without a region-specific trend term.[9] It allows for heterogeneity in the autoregressive coefficient of the Dickey-Fuller regression and a single unobserved common factor with heterogeneous factor loadings in the data. Therefore, it takes into account cross-section dependence.[10] The statistic is constructed from panel-member-specific (A)DF regressions where cross-section averages of the dependent and independent variables (including lagged differences to account for serial correlation) are included in the model (CADF regressions). Testing for the null of a unit root is based on the t-ratio of the first-order autoregressive parameter. To construct a panel statistic, the *t*-values are pooled across cross sections. A standardized version of the test is asymptotically distributed as standard normal under the joint null hypothesis of nonstationarity for all cross sections. If the null is rejected, the series is stationary at least for one panel member. Under the null of nonstationarity the test statistic has a non-standard distribution. The test is found to have good size and power properties, even when N and T are relatively small.

[9] There should be a careful use of a deterministic trend or "*…otherwise results can be misleading…*" (see, Ahking, 2002; p.51).

[10] Tests which do not account for dependence, when is exists, suffer from huge size distortions, which increase with the number of cross sections (Banerjee et al., 2004; 2005).

For implementation of the panel unit root tests, we use the Bartlett kernel. All bandwidths and lag lengths equal $4\left(T/100\right)^{\frac{2}{9}} \approx 2.89$, where T = 23 in our case (Basher and Westerlund, 2007). So, the maximum lag length lies between 2 and 3. Too few lags reduce the size of a unit root test, while too many lags reduce its power (Campbell and Perron, 1991). We conduct the panel unit root tests with the lag length equal to 2, following Martins (2011) and Jaunky (2012). For comparison, we also present the findings from the Maddala and Wu (1999) test, which allows for heterogeneity in the autoregressive coefficient of the Dickey-Fuller regression, but ignores cross-section dependence.

Table 2 presents the results. Employed tertiary education graduates and public capital stock per worker are I(2). Employment density is stationary in levels and GDP per worker, primary and secondary education graduates, labor force density and population concentration are stationary in differences. Given these findings, we can not study the long-run relationship among levels of the variables. So, we do not conduct panel cointegration tests and proceed directly to the estimation of the short-run relationship among growth rates of the variables with a methodology described below, which is robust to non-cointegration (Bronzini and Piselli, 2009).

[INSERT TABLE 2 HERE]

4.4 Estimation methodology

Having established the order of integration of our variables, we proceed with the estimation of Equation (9) in first differences, using panel econometric techniques. This way we mitigate endogeneity, since education is highly income-elastic and the Greek economy, like other high-income economies, is dominated by the service sector which requires a well-educated workforce (Catao and Solomou, 2005; Catao and Terrones, 2005). Moreover, we are able to estimate the short-run effects of education, public capital and density on labor productivity. We allow for heterogeneity in the relationship between education and productivity across regions by including region-specific fixed effects, input parameters and factor loadings.

There are alternative procedures for estimating Equation (9). We apply various estimators, which incorporate different assumptions about the underlying data generating process.

Generally, the simple pooled estimators assume a fully homogeneous coefficient model in which all slope and intercept parameters are identical across regions, meaning that regions follow the same underlying model relating productivity to the right-hand side variables. However, from the work of Durlauf and Johnson (1995), Lee et al. (1998) and Temple (1999a,b) among others, we know that this is not a trivial assumption, so allowing parameter heterogeneity can change results of growth regressions in very important ways. For instance, errors are non-stationary if 'true' technology parameters are heterogeneous and input variables are non-stationary.

In light of these, we initially use the Fixed-effect (FE hereafter) estimator, which allows only the intercepts to differ across regions as a benchmark. In other words, we assume common parameters on factor inputs and convergence rates and heterogeneity with respect to TFP growth across regions. We ignore cross-section dependence in the form of unobserved common factors. However, our estimator is robust to cross-sectional heteroscedasticity and within panel serial correlation. Additionally, we employ the Arellano and Bover (1995) - Blundell and Bond (1998) (AB-BB hereafter) estimator. Blundell and Bond (1998), show that the lagged level instruments in the Arellano and Bond (1991) estimator[11] become weak as the autoregressive process becomes too persistent or the ratio of the variance of the cross-section effects to the variance of the idiosyncratic error becomes too large and proposed a system GMM estimator, which we use, building on Arellano and Bover (1995). The latter estimator employs moment conditions in which lagged differences are instruments for the level equation, in addition to the moment conditions of lagged levels as instruments for the differenced equation. We treat the explanatory variables as endogenous and account for heteroscedasticity in the data-generating process. This estimator still assumes common factor input parameters and common impact of unobservables, although it solves the identification problem due to the correlation between inputs and unobservables (TFP in our case) (see discussion in section 4.2).

For these reasons, we finally implement the Augmented Mean Group (AMG) estimator (Eberhardt and Teal, 2009, 2012a, 2012b), which allows for cross-section dependence.

[11] This estimator requires first differencing and lags of the dependent and explanatory variables as instruments (Caselli et al., 1996). First differencing removes region-specific effects, which are a potential source of omitted variable bias and deals with series non-stationarity.

Also, it does not require pre-testing for cointegration. It was developed with production function estimation in mind, where unobservables represent TFP. We choose AMG to Common Correlated Effects Mean Group (CCEMG hereafter) estimator (Pesaran, 2006), because the latter treats f_t as a nuisance, which must be accounted for, but is not of particular interest for the empirical analysis. But in our empirical model, we view TFP as a 'measure of our ignorance' (Abramowitz, 1956), reflecting a wide set of factors which can shift the production possibility frontier (for instance "…resource endowments, climate, institutions, and so on…", Mankiw et al., 1992; pp.410-411).[12]

The AMG estimator first involves estimation of a pooled regression model with period dummies by first difference OLS and collection of the coefficients on the (differenced) dummies, which correspond to the estimated cross-region average of TFP evolution. This is called the "common dynamic process". Second, the group-specific regression model is augmented with this TFP process either: (a) as an explicit variable or (b) imposes on each group member a unit coefficient by subtracting the estimated process from the dependent variable. The regression model includes an intercept, which corresponds to time-invariant fixed effects (TFP levels). Third, the region-specific parameters are averaged across the panel. Therefore, AMG allows for heterogeneous technology parameters and heterogeneous factor loadings. The standard errors reported in the averaged regression results are constructed following Pesaran and Smith (1995) and test the significance of the average coefficients.

5. Estimation results

We estimate Equation (9) in first differences using up to 2 lags of the variables.[13] We do this, because: i) we are interested in the dynamic effects of education on productivity growth, and ii) variables are not cointegrated (Bronzini and Piselli, 2009). Here we should note that we employ lags also for the density variables, because technological agglomeration externalities take time to materialize, since they depend on the formation of communication networks which can not take place quickly (Brülhart and Mathys, 2008; Henderson, 1997). We also test the common factor restrictions and fail to accept them, so

[12] In Monte Carlo simulations, the AMG performed similarly well as the CCEMG in panels with nonstationary variables (cointegrated or not) with common factor errors (Eberhardt and Bond, 2009).

[13] According to Loayza and Ranciere (2006) when the emphasis is on the short-run parameters, it is recommended to impose a common lag structure across regions. In order to preserve the degrees of freedom while allowing for reasonably rich dynamics, we set the lag length equal to 2.

we estimate an unrestricted dynamic panel model. We report the results of FE, AB-BB and AMG estimation for completeness, but we emphasize evidence from the AMG estimator for reasons outlined in the previous section.[14]

Overall, three models are estimated. First, following our theoretical model (see Equation 5), the relationship between labor productivity growth and education at all levels, including public capital and employment density as control variables, is estimated. In the AMG estimations we impose a common dynamic process with a unit coefficient for TFP, subtracting the estimated process from the dependent variable, after having tested this assumption and verified it. Results in Table 3 indicate that primary education does not play a role in productivity growth. On the contrary, a 1 standard deviation increase in growth of the share of employees with secondary education reduces productivity growth around 2% contemporaneously, 2.6% after one year and 3% after two years. In the same manner, if growth of employees with tertiary education rises by 1 standard deviation, productivity growth is enhanced by 2.5% and 2.9% contemporaneously and with a one-year lag, respectively.

[INSERT TABLE 3 HERE]

This evidence on primary and secondary education could be explained by the low skills of Greek graduates compared to corresponding OECD graduates, as indicated by student performance in international standardized tests, such as PISA (OECD, 2010).[15] Several problems have been identified in primary and secondary education in Greece affecting human capital. These include, among others, excessively small student-teacher ratios and class sizes, low teacher salaries, lack of external assessment and evaluation of schools, teachers, students and the education system as a whole (OECD, 2011). Additional weaknesses constitute extremely centralized governance of the education system, limited opportunities for professional development of education personnel, fragmented budgeting procedures, limited accountability over outcomes and limited school competition. Our

[14] We have examined our dataset for possible multicollinearity and found no evidence of such a problem with the exception of the high correlation between employment density and public capital. However, the fact that results do not change when we replace employment density with labor force density and population concentration implies that the above high correlation does not distort our findings in any important way.

[15] The Programme for International Student Assessment (PISA) is an international study conducted by the OECD in its member and non-member nations and examines 15-year-old school pupils' scholastic performance on mathematics, science, and reading. It was first performed in 2000 and is repeated every three years.

findings are only partially in line with Asteriou and Agiomirgianakis (2001), who find a positive relationship between enrollments in primary, secondary, tertiary education and GDP in Greece. However, we believe that our human capital measures reflect more accurately the skills of the working population, since for example some students enrolled in secondary school may not finish it or even if they complete it, they may not be working for some time. Concerning primary and tertiary education, we obtain the same findings with Ramos et al. (2010) who examine Spanish regions at the same level of disaggregation. As for tertiary education, we reach the same conclusions with Arbia et al. (2010), Cuaresma et al., (2012), and Cuaresma and Feldkircher (2012) for EU regions and Abel and Gabe (2011) for US cities. Our evidence is also in line with that of Petrakis and Stamatakis (2002) and Vandenbussche et al. (2006), who argue that in developed economies tertiary education is more important relative to lower education levels.[16]

We also find that public investment has an insignificant growth effect. Although this is in contrast to some related literature (Munnell, 1992; Bronzini and Piselli, 2009), research on Greek regions shows mixed results (Rodríguez-Pose et al., 2012; Lambrinidis et al., 2005). A possible explanation would be the inefficient spatial and functional allocation of public investment, due to e.g. political considerations. These considerations have to do with the party in power at the central government level, the regional vote share in favor of the governing party or the difference in the regional vote shares between the ruling party and the main opposition party (Yamano and Ohkawara, 2000; Johansson, 2003; Castells and Solé-Ollé, 2005; Rodríguez-Pose et al., 2012).

Employment density is not statistically significant, implying that the negative effects of density just outweigh the positive ones in Greek regions. This is different with respect to Ciccone and Hall (1996), Ciccone (2002), Brülhart and Mathys (2008), who find that employment density boosts labor productivity in the US states and European regions. Our findings may be due to the fact that labor and product markets were more rigid in Greece compared to the most European countries and the US during the period under consideration (OECD, 2004 & 2012). Finally, we find conditional convergence, as lagged growth reduces current growth by 0.29% and 0.25% after one and two years respectively.

[16] Studies using micro data show that private returns to education in Greece are positively associated with years of education (Magoula and Psaharopoulos, 1999; Mitrakos et al., 2010).

This is in line with Christopoulos and Tsionas (2004) and Michelis et al. (2004). For comparison, in the FE case, we find conditional convergence and a negative effect of public capital contemporaneously and of primary education regarding lagged variables. We also reveal a mixed impact of secondary education and a positive effect of tertiary education after 2 years. Finally, employment density does not influence growth overall meaning that its positive and negative effects exactly offset each other. In the AB-BB estimation, none of the parameters is statistically significant. Overall, the evidence is only partially in line with that obtained by AMG estimation. However, we do not put much weight on FE and AB-BB results for reasons mentioned in Section 4.4.

Next, we repeat our estimations using the same methodology as before, but including labor force density instead of employment density, as an indicator of density-related effects (see Table 4 below). Our findings imply that secondary education graduates have a negative impact on labor productivity growth of 2.6% with a 2-year lag, similar to the impact estimated before in the same time horizon. Tertiary education graduates exert a positive contemporaneous growth impact of 2.6%, very similar to the one estimated before in the same time frame. Once again, primary education and public capital do not seem to matter for growth. As regards the density variable, growth in labor force density has no statistically significant growth influence, indicating the equally strong opposing forces of agglomeration and congestion on productivity. Finally, we confirm the tendency for convergence in terms of labor productivity growth of Greek regions with coefficients equal to -0.25 and -0.22 for the 1-year and 2-year lags, respectively, which are very similar to the ones estimated in the first specification of the model. Overall, results are analogous to the ones estimated in the previous model, both quantitatively and qualitatively, which is a sign of robustness. Regarding FE and AB-BB estimates, they are almost identical with the previous model.

[INSERT TABLE 4 HERE]

In the same spirit, we employ population concentration, i.e. relative population density, and obtain similar findings (see Table 5 below). In the AMG estimation, secondary and tertiary education growth affect productivity growth negatively and positively respectively with a 2-year lag. Primary education, public capital and population concentration do not influence productivity. Finally, there is conditional convergence of Greek regions in terms of labor productivity dynamics. Concerning FE and AB-BB estimates, findings are similar with the other two models. Turning to the diagnostics, tests show that in all cases residuals from AMG estimations do not suffer from cross-section dependence and are stationary. On the other hand, both FE and AB-BB estimators yield cross-sectionally dependent, but mostly stationary residuals.[17]

[INSERT TABLE 5 HERE]

As a synopsis, our empirics have emphasized cross-section dependence and parameter heterogeneity. In other words, we have examined the importance of spillovers, unobserved shocks and production technology heterogeneity across Greek regions. We present a number of dynamic specifications and different estimators. Our initial results – which exhibit cross-sectionally dependent residuals and are based on the common production technology assumption – show that tertiary education plays no role for labor productivity growth, while primary education has a growth retarding impact. However, once we tackle these issues, our evidence is modified pointing to the relevance of human capital in the form of tertiary education for productivity dynamics. This is an apparent signal that spillovers and technology differences are of utmost importance in growth empirics.

6. Conclusions

Most research on production function estimation assumes cross-section independence and parameter homogeneity in production technology. However, both assumptions are strong for regional data. In this paper, we construct a rich dataset and estimate production functions for Greek regions employing both standard and non-stationary panel techniques allowing for cross-section dependence as well as parameter heterogeneity. We put special emphasis on the impact of the three education levels on labor productivity. This allows us to estimate their differentiated effects on growth.

After verifying the existence of cross-section dependence, we conduct the appropriate pane unit root tests and proceed to empirical estimation of three models in growth form, focusing on short-run dynamics, which are very relevant from a policy point of view. Employing the AMG estimator, we find robust evidence of a strong positive impact of tertiary education and a weaker negative effect of secondary education on labor

[17] Arellano and Bond tests can not reject the hypothesis of no serial correlation at order two in the first-differenced errors in all three models for AB-BB estimations. Thus, there is no evidence of models' misspecification.

productivity growth. On the contrary, we do not infer such an influence of primary education, public capital and density, irrespectively of density measure used. Comparing these results with the findings from fixed effects and dynamic panel estimators, for almost all variables, evidence is qualitatively and quantitatively different. Thus, neglecting cross-section dependence and parameter heterogeneity creates bias in the estimated input elasticities. Our findings support related studies arguing that lower levels of education are important in developing countries, while tertiary education is significant for developed ones (Petrakis and Stamatakis, 2002; Papageorgiou, 2003). This is because a knowledge-based economy requires highly skilled labor force in order to exploit its full growth potential (Pereira and Aubyn, 2009).

A number of implications can be derived from our results. First, if policy makers do not account for spillovers and technology heterogeneity, regional growth policies will be misguided. Specifically, growth-enhancing policies should be directed at relatively spacious areas in Greece. Second, our evidence confirms that the expansion of tertiary education is the most effective instrument for reducing regional disparities in terms of labor productivity. The higher education – productivity link may be materialized through various channels, e.g. higher innovative capacity and increased capability for adoption of advanced foreign technology. This line of research can further be enhanced by the improvement of human capital measurement, taking more comprehensive account of its quality and other forms of education (job training) when data become available. Moreover, a thorough quantification of the spillovers, e.g. using spatial econometrics, might be useful. These are left for future work.

References

Abel, J.R, Dey, I., Gabe, T.M., 2012. Productivity and the Density of Human Capital, Journal of Regional Science, 52, 4, pp.562-586.

Abel, J.R., Gabe, T.M., 2011. Human Capital and Economic Activity in Urban America, Regional Studies, 45, 8, pp.1079–1090.

Abramowitz, M., 1956. Resource and output trends in the United States since 1870. American Economic Review 46, 2, pp.5–23.

Ahking F.W., 2002. Model mis-specification and Johansen's co-integration analysis: an application to the US money demand. Journal of Macroeconomics, 24, 1, pp.51–66.

Andersson, F., Burgess, S., Lane, J.I., 2007. Cities, Matching and the Productivity Gains of Agglomeration. Journal of Urban Economics, 61, pp.112–118.

Arbia, G., Battisti, M., Di Vaio, G., 2010. Institutions and geography: Empirical test of spatial growth models for European regions. Economic Modelling 27, pp.12–21.

Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies, 58, pp.277-297.

Arellano, M., Bover, O., 1995. Another look at the instrumental variable estimation of error-components models. Journal of Econometrics, 68, pp.29-51.

Asteriou, D., Agiomirgianakis, G.M. 2001. Human capital and economic growth: Time series evidence from Greece. Journal of Policy Modeling, 23, pp.481–489.

Azariadis, C., Drazen, A., 1990. Threshold externalities in economic development. Quarterly Journal of Economics, 105, pp.501–26.

Baltagi BH, 2008. Econometric analysis of panel data, 4th edition Wiley, Chichester.

Banerjee, A., Marcellino, M., Osbat, C., 2004. Some Cautions on the Use of Panel Methods for Integrated Series of Macroeconomic Data. Econometrics Journal, 7, pp.322–40.

Banerjee, A., Marcellino, M., Osbat, C., 2005. Testing for PPP: Should We Use Panel Methods? Empirical Economics, 30, 1, pp.77–91.

Barro, R.J., Lee, J.W., 2010. A new data set of educational attainment in the world, 1950–2010. NBER Working Paper No. 15902. (Educational Attainment Dataset, Version 1.2)

Barro, R.J., Sala-i-Martin, X., 2004. Economic growth. Cambridge, MA: MIT Press.

Basher S.A, Westerlund, J., 2007. Is there really a unit root in the inflation rate? More evidence from panel data models. Applied Economics Letters 15, 3, pp.161–164.

Blundell, R., Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics, 87, pp.115-143.

Bronzini, R., Piselli, P., 2009. Determinants of long-run regional productivity with geographical spillovers: The role of R&D, human capital and public infrastructure. Regional Science and Urban Economics, 39, pp.187–199.

Brülhart, M., Mathys, N.A., 2008. Sectoral agglomeration economies in a panel of European regions. Regional Science and Urban Economics, 38, pp.348–362.

Campbell, J.Y., Perron, P., 1991. Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots. NBER Chapters, in: NBER Macroeconomics Annual 1991, 6, National Bureau of Economic Research, Inc., pp.141–220

Caselli, F., Esquivel, G., Lefort, F, 1996. Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics, Journal of Economic Growth, 1, 3, pp.363-89.

Castells, A., Solé-Ollé, A. 2005. The regional allocation of infrastructure investment: The role of equity, efficiency and political factors. European Economic Review, 49, pp.1165-1205.

Catao, L.A.V., Solomou, S.N., 2005. Effective exchange rates and the classical gold standard adjustment. American Economic Review, 95, pp.1259–1275.

Catao, L.A.V., Terrones, M.E., 2005. Fiscal deficits and inflation. Journal of Monetary Economics, 52, pp.529–554.

Cheshire, P.C. Malecki, E.J., 2004. Growth, development, and innovation: a look backward and forward. Papers in Regional Science, 83, pp.249-267.

Christopoulos D., Tsionas E, 2004. Convergence and regional productivity differences: evidence from Greek prefectures. Annals of Regional Science, 38, pp.387–396.

Chudik, A., Pesaran, M.H., Tosetti, E. 2011. Weak and strong cross-section dependence and estimation of large panels. Econometrics Journal, 14,1, C45-C90.

Ciccone, A., 2002. Agglomeration effects in Europe. European Economic Review 46 (2), pp.213–227.

Ciccone, A., Hall, R., 1996. Productivity and the density of economic activity. American Economic Review 86, 1, 54–70.

Cohen, J.P., Morrison Paul, C.J., 2009. Agglomeration, Productivity, and Regional Growth: Production Theory Approaches, in: R. Capello and P. Nijkamp (Eds.), Handbook of Regional Growth and Development Theories. Cheltenham: Edward Elgar.

Cuaresma, J.C., Doppelhofer, G. Feldkircher, M., 2012: The Determinants of Economic Growth in European Regions, Regional Studies, DOI:10.1080/00343404.2012.678824 (forthcoming)

Cuaresma, J.C., Feldkircher, M., 2012. Spatial filtering, model uncertainty and the speed of income convergence in Europe, Journal of Applied Econometrics, DOI: 10.1002/jae.2277

Di Liberto, A., 2007. Education and Italian regional development. Economics of Education Review, 27(1), pp.94–107.

Ding, S., Knight, J., 2011. Why has China Grown So Fast? The Role of Physical and Human Capital Formation Oxford Bulletin of Economics and Statistics, 73, 2, pp.141-174.

Durlauf, S.N., Kourtellos, A., Minkin, A., 2001. The local Solow growth model. European Economic Review 45(4–6), pp.928–940.

Durlauf, S., Johnson, P. 1995. Multiple regimes and cross-country growth behaviour. Journal of Applied Econometrics, 10, pp.365–84.

Eberhardt, M., Teal, F., 2011. Econometrics for Grumblers: A New Look at the Literature on Cross-Country Growth Empirics, Journal of Economic Surveys, 25, 1, pp.109–155.

Eberhardt, M., Teal, F., 2012a. Productivity analysis in global manufacturing production, mimeo, March, pp.1-32.

Eberhardt, M., 2012b. Estimating panel time series models with heterogeneous slopes, Stata Journal, 12, 1, pp.61-71.

Eberhardt, M., Bond, S., 2009. Cross-section dependence in nonstationary panel models: A novel estimator. MPRA Paper 17870, University Library of Munich.

Granger, C.W.J., 1988. Aggregation of time series variables: a survey. Federal Reserve Bank of Minneapolis Discussion Paper/Institute for Empirical Macroeconomics #1.

Granger, C.W.J., Siklos, P.L., 1995. Systematic sampling, temporal aggregation, seasonal adjustment, and cointegration theory and evidence. Journal of Econometrics 66, 1–2, pp.357–369.

Griliches, Z., 1997. Education, Human capital and growth: a personal perspective, Journal of Labor Economics, 5, 1, pp.330–344.

Hamming, R.W. 1973. Numerical methods for scientists and engineers. New York: McGraw-Hill, pp.349-356.

Henderson, V.J., 1997. Externalities and industrial development. Journal of Urban Economics, 42, 3, 449-470.

Jaunky, V.C., 2012. Democracy and economic growth in Sub-Saharan Africa: a panel data approach, Empirical Economics, DOI 10.1007/s00181-012-0633-x

Johansson, E. 2003. Intergovernmental grants as a tactical instrument: empirical evidence from Swedish municipalities. Journal of Public Economics 87, 5-6, pp.883-915.

Kapetanios, G., Pesaran, M.H., Yamagata, T., 2011. Panels with non-stationary multifactor error structures," Journal of Econometrics, 160, 2, pp.326-348.

Karnik, A Lalvani, M., 2012. Growth performance of Indian states, Empirical Economics, 42(1), 235-259.

Kosfeld, R., Lauridsen, J., 2004. Dynamic spatial modelling of regional convergence processes. Empirical Economics, 29, pp.705-722.

Krueger, A.B., Lindahl, M. 2001. Education for growth: why and for whom? Journal of Economic Literature, 39(4), pp.1101-36.

Lambrinidis, M., Psycharris, Y., Rovolis, A., 2005. Regional allocation of public infrastructure investment: the case of Greece. Regional Studies, 39(9), pp.1231-44.

Lee, K., Pesaran, M. H., Smith, R., 1998. Growth empirics: a panel data approach – a comment. Quarterly Journal of Economics, 113(1), pp.319-23.

Lee, K., Pesaran, M.H., Smith, R., 1997. Growth and convergence in a multi-country empirical stochastic Solow model. Journal of Applied Econometrics 12, 4, pp.357-392.

Loayza, N., Ranciere, R., 2006. Financial development, financial fragility, and growth. Journal of Money, Credit and Banking, 38, pp.1051-1076.

Louri, H., 1989. Regional policy and investment behaviour: the case of Greece, 1971–1982. Regional Studies, 23(3), pp.231-9.

Lucas, R.E., 1988. On the mechanics of economic development. Journal of Monetary Economics, 22(1), pp.3-42.

Maddala, G., Wu, S., 1999. A comparative study of unit root tests and a new simple test. Oxford Bulletin of Economics and Statistics, 61, pp.631-652.

Magoula, T., Psacharopoulos, G., 1999. Schooling and monetary rewards in Greece: an over-education false alarm. Applied Economics, 31, pp.1589-97.

Mankiw, N., Romer, D., Weil, D. 1992. A contribution to the empirics of economic growth. Quarterly Journal of Economics, 107, 2, pp.407-37.

Martins, P.M.G., 2011. Aid absorption and spending in Africa: a panel cointegration approach. The Journal of Development Studies, 47, 12, pp.1925-1953.

McDonald, J., McMillen, D. 2007. Urban Economics and Real Estate: Theory and Policy. Blackwell Press.

Michelis, L., Papadopoulos, A.P, Papanikos, G.T. 2004. Regional Convergence in Greece in the 1980s: An Econometric Investigation. Applied Economics 36, pp.881-888.

Mitrakos, T., Tsakloglou, P., Holezas, I. 2010. Wage determinants in Greece with emphasis on tertiary education graduates. Bank of Greece Economic Bulletin, No.34 (in Greek).

Munnell, A. H., 1992. Infrastructure investment and economic Growth. Journal of Economic Perspectives, 6, pp.189-198.

Nelson, R.R., Phelps, S.E., 1966. Investment in humans, technological diffusion, and economic growth. American Economic Review, 56(1-2), pp.69-75.

Nelson, C.R., Plosser, C., 1982. Trends and random walks in macroeconomic time series: some evidence and implications. Journal of Monetary Economics 10, 2, pp.139-162.

OECD, 2004. OECD Employment Outlook - Chapter 2 Employment Protection Regulation and Labour Market Performance, OECD Publishing.

OECD, 2011. Education Policy Advice for Greece, Strong Performers and Successful Reformers in Education, OECD Publishing. (http://dx.doi.org/10.1787/9789264119581-en; last access June 20,2013)

OECD, 2010. PISA 2009 at a Glance, OECD Publishing. (http://dx.doi.org/10.1787/9789264095298-en; last access September 3, 2013)

Ohmae, K., 1995. The end of the nation state. Free Press, New York.

Papageorgiou, C., 2003. Distinguishing between the effects of primary and post-primary education on economic growth. Review of Development Economics, 7(4), pp.622–635.

Pablo-Romero, M. del P., Gómez-Calero, M. de la P., 2013. A translog production function for the Spanish provinces: Impact of the human and physical capital in economic growth. Economic Modelling 32, pp.77–87

Pedroni, P.,2007. Social capital, barriers to production and capital shares: implications for the importance of parameter heterogeneity from a nonstationary panel approach. Journal of Applied Econometrics, 22, 2, pp.429-451.

Pereira, J., Aubyn, M. St., 2009. What level of education matters most for growth? Evidence from Portugal. Economics of Education Review, 28(1), pp.67–73.

Pesaran, M.H., Smith, R.J. (1995). Estimating long-run relationships from dynamic heterogeneous panels. Journal of Econometrics, 68, pp.79-113.

Pesaran, M.H., 2004. General Diagnostic Tests for Cross Section Dependence in Panels, Working Paper 435, University of Cambridge, Cambridge.

Pesaran, M., 2007. A Simple Panel Unit Root Test in the Presence of Cross Section Dependence. Journal of Applied Econometrics, 22, pp.265-312.

Petrakis, P.E., Stamatakis, D., 2002. Growth and educational levels: A comparative analysis. Economics of Education Review, 21(5), pp.513-521.

Phillips, P.C.B., Sul, D., 2003 Dynamic panel estimation and homogeneity testing under cross section dependence. Econometrics Journal, 6, 1, pp.217-259.

Piras, G., Postiglione, P., Aroca, P., 2012. Specialization, R&D and productivity growth: evidence from EU regions. Annals of Regional Science, 49, 1, pp. 35-51.

Press, W.H., Teukolsky, S.A, Vetterling, W.T., Flannery, B.P. 2007. Numerical recipes: the art of scientific computing. Cambridge: Cambridge University Press. pp.120-124.

Pritchett, L., 2001. Where has all the education gone? World Bank Economic Review, 15, 3, pp.367-91.

Ramos, R., Suriñach, J., Artís, M., 2010. Human capital spillovers, productivity and regional convergence in Spain. Papers in Regional Science, 89, 2, pp.435-447.

Rodríguez-Pose, A., Psycharis, Y., Tselios, Y., Public investment and regional growth and convergence: evidence from Greece, CEPR DP 9011, 6/12.

Romer, P.M., 1990. Endogenous technological change. Journal of Political Economy, 98, 5, pp.71-102.

Roos, 2005, M.W.M. How important is geography for agglomeration? Journal of Economic Geography, 5, pp.605–620.

Rovolis, A., Spence, N., 2002. Promoting Regional Economic Growth in Greece by Investing in the Public Infrastructure" Environment and Planning C: Government and Policy, 20, 3, pp.393-419.

Sari, R., Soytas, U., 2006. Income and education in Turkey: A multivariate analysis. Education Economics, 14, 2, pp.181-196.

Soukiazis, E., Antunes, M., 2011. Is foreign trade important for regional growth? Empirical evidence from Portugal. Economic Modelling, 28, pp.1363–1373.

Stata Reference Manual, Release 11, volume D, 2009, Stata Press, College Station, Texas.

Storper, M. 1997. The regional world: territorial development in a global economy. Guilford Press, New York.

Temple, J., 1999a. A positive effect of human capital on growth. Economics Letters, 65, 1, pp.131-134.

Temple, J., 1999b. The new growth evidence. Journal of Economic Literature, 37(1), pp.112–56.

United Nations 2000, General Assembly, 18 September (A/RES/55/2), Fifty-fifth session, Agenda item 60 (b), 00 55951, Resolution adopted by the General Assembly [without reference to a Main Committee (A/55/L.2)], 55/2. United Nations Millennium Declaration

Vandenbussche, J., Aghion, P., Meghir, C., 2006. Growth, distance to frontier and composition of human capital. Journal of Economic Growth, 11, 2, pp.97-127.

Yamano, N., Ohkawara, T., 2000. The regional allocation of public investment: efficiency or equity? Journal of Regional Science, 40, 2, pp.205-229.

TABLE 1: CROSS-SECTION CORRELATION

PANEL A: LEVELS

	Ln GDP	*Ln* Public Capital	*Ln* Primary Graduates	*Ln* Secondary Graduates	*Ln* Tertiary Graduates	*Ln* Employment Density	*Ln* Labor Force Density	*Ln* Population Concentration
avg ρ	0.606	0.334	0.947	0.995	0.985	0.334	0.388	1.000
avg $\lvert\rho\rvert$	0.493	0.683	0.947	0.995	0.985	0.681	0.855	1.000
CD	84.40	57.26	162.11	170.39	168.63	57.13	66.43	171.25
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

PANEL B: FIRST DIFFERENCES

	Δ*Ln* GDP	Δ*Ln* Public Capital	Δ*Ln* Primary Graduates	Δ*Ln* Secondary Graduates	Δ*Ln* Tertiary Graduates	Δ*Ln* Employment Density	Δ*Ln* Labor Force Density	Δ*Ln* Population Concentration
avg ρ	0.156	0.574	0.649	0.796	0.657	0.483	0.180	1.000
avg $\lvert\rho\rvert$	0.273	0.992	0.649	0.796	0.678	0.754	0.569	1.000
CD	26.20	96.21	108.69	133.29	109.97	80.96	30.16	167.48
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: We present the average and average absolute correlation coefficients across the $N(N - 1)$ sets of correlations. CD reports the Pesaran (2004) cross-section dependence statistic, which is distributed $N(0, 1)$ under the null of cross-section independence. Panels A and B test the variable series in levels and first differences, respectively.

TABLE 2: PANEL UNIT ROOT TESTS

		GDP		Primary Graduates		Secondary Graduates	
		without trend	*with trend*	*without trend*	*with trend*	*without trend*	*with trend*
Maddala & Wu (1999)	*Level*	55.01 (0.98)	73.59 (0.98)	1.70 (1.00)	28.81 (1.00)	155.21 (0.01)	39.73 (1.00)
	First Difference	299.29 (0.00)	207.50 (0.00)	129.51 (0.03)	359.71 (1053)	220.12 (0.03)	220.45 (0.00)
Pesaran (2007)	*Level*	-1.30 (0.09)	2.40 (0.99)	8.18 (1.00)	4.78 (976)	1.06 (0.85)	7.08 (1.00)
	First Difference	-5.40 (0.00)	-2.92 (0.00)	-3.40 (0.00)	-2.73 (0.00)	-2.24 (0.01)	-3.65 (0.00)
		Tertiary Graduates		**Public Capital**		**Employment Density**	
		without trend	*with trend*	*without trend*	*with trend*	*without trend*	*with trend*
Maddala & Wu (1999)	*Level*	229.03 (0.00)	37.13 (1.00)	202.56 (0.00)	59.36 (1.00)	242.86 (0.00)	202.42 (0.00)
	First Difference	93.41 (0.71)	151.83 (0.01)	201.25 (0.00)	256.32 (0.00)	74.44 (0.98)	114.43 (0.18)
Pesaran (2007)	*Level*	-1.30 (0.09)	10.28 (1.00)	11.93 (1.00)	12.26 (1.00)	8.94 (1.00)	-6.81 (0.00)
	First Difference	5.24 (1.00)	7.48 (1.00)	11.79 (1.00)	-9.07 (1.00)	-12.42 (0.00)	-9.07 (0.00)
		Labor Force Density		**Population Concentration**			
		without trend	*with trend*	*without trend*	*with trend*		
Maddala & Wu (1999)	*Level*	244.39 (0.00)	216.97 (0.00)	8.62 (1.00)	50.73 (1.00)		
	First Difference	148.61 (0.01)	171.30 (0.00)	155.40 (0.01)	60.89 (1.00)		
Pesaran (2007)	*Level*	8.44 (1.00)	-2.28 (0.01)	32.10 (1.00)	30.79 (1.00)		
	First Difference	0.45 (0.67)	-5.02 (0.00)	32.10 (1.00)	30.79 (1.00)		

Notes: we report the Fisher statistic and associated p-value the Maddala and Wu (1999) test and the standardised Z-tbar statistic and the corresponding p-value for the Pesaran (2007) test. For both tests the null hypothesis is that all series are nonstationary.

TABLE 3: ESTIMATION RESULTS (WITH EMPLOYMENT DENSITY)

	FE estimates		AB-BB estimates		AMG estimates	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Δ*Ln* Primary Graduates	-0.19 **	0.08	-0.24	0.23	-0.29	0.57
Δ*Ln* Primary Graduates (-1)	-0.18 **	0.72	-0.22	0.20	0.30	0.54
Δ*Ln* Primary Graduates (-2)	-0.28 ***	0.71	-0.20	0.21	0.06	0.41
Δ*Ln* Secondary Graduates	0.17 **	0.78	0.16	0.21	-0.47 **	0.24
Δ*Ln* Secondary Graduates (-1)	0.21 ***	0.08	0.21	0.24	-0.63 **	0.31
Δ*Ln* Secondary Graduates (-2)	-0.25 ***	0.07	-0.20	0.14	-0.73 ***	3.59
Δ*Ln* Tertiary Graduates	-0.03	0.13	-0.04	0.70	1.38 *	0.76
Δ*Ln* Tertiary Graduates (-1)	-0.14	0.14	-0.25	0.76	1.59 **	0.91
Δ*Ln* Tertiary Graduates (-2)	0.31	0.18	0.30	0.87	0.73	0.89
Δ*Ln* Public Capital	-16.79 *	9.63	-11.65	73.40	85.26	118.39
Δ*Ln* Public Capital (-1)	24.50	19.28	25.57	13.49	20.49	156.04
Δ*Ln* Public Capital (-2)	-9.63	9.99	-10.55	70.47	-26.71	111.75
Δ*Ln* Employment Density	-0.23	0.95	-2.50	5.20	1.48	1.93
Δ*Ln* Employment Density (-1)	1.15	0.73	-1.31	5.37	-0.36	2.41
Δ*Ln* Employment Density (-2)	0.24	0.72	-0.47	2.96	1.67	2.23
Δ*Ln* GDP (-1)	-0.16 ***	0.03	-0.15	0.93	-0.29 ***	0.08
Δ*Ln* GDP (-2)	-0.09 **	0.03	-0.56 ***	0.86	-0.25 ***	0.08
Constant	0.01	0.01	0.0	0.72	0.32	0.09
Order of Integration	I(0)		I(0)		I(0)	
AB Test AR (1)	-		-4.23 (0.00)		-	
AB Test AR (2)	-		-0.13 (0.89)		-	
CD Test	73.89 (0.00)		67.35 (0.00)		0.84 (0.40)	
Observations	1020		1020		1020	

Notes: Dependent variable: *ΔlnGDPpw*. All the variables in natural log differences. Time Period: 1981-2003. 51 regions. ***Coefficient Significant at 1%; **Coefficient Significant at 5%; *Coefficient Significant at 10%. FE: Fixed Effects estimator with region-specific effects and cluster-robust standard errors; AB-BB: two-step system GMM estimator with robust standard errors and all variables considered endogenous; AMG: Augmented Mean Group estimator. *Diagnostics*: The order of integration of the residuals is determined using the Pesaran (2007) CIPS test for H_o of nonstationarity (full results available upon request); AB tests: Arellano and Bond tests for 1st and 2nd order serial correlation in the first-differenced errors with H_o of no serial correlation, z-statistics are provided with p-values in parentheses; CD test: Pesaran (2004) test for H_o of cross-sectionally independent errors, test statistics are provided with p-values in parentheses.

TABLE 4: ESTIMATION RESULTS (WITH LABOR FORCE DENSITY)

	FE estimates		AB-BB estimates		AMG estimates	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Δ*Ln* Primary Graduates	-0.19 *	0.08	-0.25	0.61	-0.78	0.72
Δ*Ln* Primary Graduates (-1)	-0.18 **	0.07	-0.26	0.47	0.58	0.44
Δ*Ln* Primary Graduates (-2)	-0.29 ***	0.07	-0.26	0.38	-0.24	0.41
Δ*Ln* Secondary Graduates	0.17 **	0.79	0.21	0.26	-0.25	0.23
Δ*Ln* Secondary Graduates (-1)	0.21 ***	0.08	0.27	0.21	-0.27	0.27
Δ*Ln* Secondary Graduates (-2)	-0.25 ***	0.07	-0.17	0.18	-0.61 ***	0.22
Δ*Ln* Tertiary Graduates	-0.02	0.14	-0.21	1.21	1.47 **	0.79
Δ*Ln* Tertiary Graduates (-1)	-0.13	0.14	-0.36	1.77	0.72 *	0.74
Δ*Ln* Tertiary Graduates (-2)	0.29	0.20	0.38	1.88	0.63	0.73
Δ*Ln* Public Capital	-17.91 *	9.89	-14.57	71.86	116.02	71.67
Δ*Ln* Public Capital (-1)	27.39	19.62	25.15	13.01	-182.64	127.93
Δ*Ln* Public Capital (-2)	-10.27	10.30	-12.02	73.13	-117.41	83.18
Δ*Ln* Labor Force Density	0.01	0.01	0.01	0.01	0.06	0.14
Δ*Ln* Labor Force Density (-1)	0.01	0.01	-0.01	0.01	-0.01	0.15
Δ*Ln* Labor Force Density (-2)	0.01	0.01	-0.01	0.01	0.22	0.15
Δ*Ln* GDP (-1)	-0.17 ***	0.03	-0.13	0.49	-0.29 ***	0.08
Δ*Ln* GDP (-2)	-0.09 **	0.03	-0.06 ***	0.38	-0.22 ***	0.07
Constant	0.01	0.01	0.01	0.19	-0.13	0.12
Order of Integration	I(0)		I(0)		I(0)	
AB Test AR (1)	-		-1.38 (0.16)		-	
AB Test AR (2)	-		-0.19 (0.84)		-	
CD Test	76.22 (0.00)		66.29 (0.00)		1.08 (0.28)	
Observations	1020		1020		1020	

Notes: Dependent variable: *ΔlnGDPpw*. All the variables in natural log differences. Time Period: 1981-2003. 51 regions. ***Coefficient Significant at 1%; **Coefficient Significant at 5%; *Coefficient Significant at 10%. FE: Fixed Effects estimator with region-specific effects and cluster-robust standard errors; AB-BB: two-step system GMM estimator with robust standard errors and all variables considered endogenous; AMG: Augmented Mean Group estimator. *Diagnostics*: The order of integration of the residuals is determined using the Pesaran (2007) CIPS test for H_o of nonstationarity (full results available upon request); AB tests: Arellano and Bond tests for 1st and 2nd order serial correlation in the first-differenced errors with H_o of no serial correlation, z-statistics are provided with p-values in parentheses; CD test: Pesaran (2004) test for H_o of cross-sectionally independent errors, test statistics are provided with p-values in parentheses.

TABLE 5: ESTIMATION RESULTS (WITH POPULATION CONCENTRATION)

	FE estimates		AB-BB estimates		AMG estimates	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Δ*Ln* Primary Graduates	-0.14 **	0.08	-0.23	0.17	-112.44	122.47
Δ*Ln* Primary Graduates (-1)	-0.14 **	0.07	-0.17	0.13	157.86	105.48
Δ*Ln* Primary Graduates (-2)	-0.27 ***	0.07	-0.22	0.15	-23.46	147.82
Δ*Ln* Secondary Graduates	0.06	0.09	0.02	0.14	-0.66	0.58
Δ*Ln* Secondary Graduates (-1)	0.17 **	0.08	0.27	0.17	-0.16	0.50
Δ*Ln* Secondary Graduates (-2)	-0.18 **	0.07	-0.14	0.14	-0.84 **	0.45
Δ*Ln* Tertiary Graduates	-0.06	0.14	-0.14	0.42	0.21	1.31
Δ*Ln* Tertiary Graduates (-1)	-0.21	0.15	-0.07	0.51	0.37	1.75
Δ*Ln* Tertiary Graduates (-2)	0.33	0.21	0.59	0.44	2.78 *	1.63
Δ*Ln* Public Capital	-18.45 *	10.08	24.98	60.85	27.66	167.02
Δ*Ln* Public Capital (-1)	32.42	20.07	-50.49	14.10	-151.28	237.07
Δ*Ln* Public Capital (-2)	-14.82	10.57	26.16	55.70	148.54	184.58
Δ*Ln* Population Concentration	0.69	0.52	-1.11	2.80	-0.53	2.41
Δ*Ln* Population Concentration (-1)	-2.85 ***	0.78	-5.58	3.46	-4.50	3.86
Δ*Ln* Population Concentration (-2)	-2.48 ***	0.61	-1.77	2.81	-0.55	3.22
Δ*Ln* GDP (-1)	-0.18 ***	0.03	-0.15 *	0.06	-0.25 ***	0.10
Δ*Ln* GDP (-2)	-0.09 **	0.03	-0.05	0.05	-0.16 **	0.08
Constant	0.02 ***	0.01	0.01	0.04	0.57	0.45
Order of Integration	I(0)		I(0)		I(0)	
AB Test AR (1)	-		-4.44 (0.00)		-	
AB Test AR (2)	-		-0.05 (0.95)		-	
CD Test	100.79 (0.00)		1.41 (0.00)		-0.72 (0.47)	
Observations	1020		1020		1020	

Notes: Dependent variable: *ΔlnGDPpw*. All the variables in natural log differences. Time Period: 1981-2003. 51 regions. ***Coefficient Significant at 1%; **Coefficient Significant at 5%; *Coefficient Significant at 10%. FE: Fixed Effects estimator with region-specific effects and cluster-robust standard errors; AB-BB: two-step system GMM estimator with robust standard errors and all variables considered endogenous; AMG: Augmented Mean Group estimator. *Diagnostics*: The order of integration of the residuals is determined using the Pesaran (2007) CIPS test for H_o of nonstationarity (full results available upon request); AB tests: Arellano and Bond tests for 1st and 2nd order serial correlation in the first-differenced errors with H_o of no serial correlation, z-statistics are provided with p-values in parentheses; CD test: Pesaran (2004) test for H_o of cross-sectionally independent errors, test statistics are provided with p-values in parentheses.

APPENDIX: DATA DEFINITION, DESCRIPTIVE STATISTICS AND CONSTRUCTION OF EDUCATION DATASET

TABLE 1: DEFINITION OF VARIABLES

VARIABLES	DESCRIPTION	SOURCE
GDP per worker	GDP per worker; in Euros, at 2000 constant prices;	Hellenic Statistical Authority, Quarterly Regional & Satellite Accounts Section
Public Capital per worker	Public capital is calculated using the investment of the state budget at regional level, applying the perpetual investment method (Rovolis & Spence, 2002). The variable is constructed by extracting payments for services rendered from total regional state budget expenditures. Per worker; in Euros, at 2000 constant prices.	Hellenic Statistical Authority, Public Sector Survey Section & authors' calculations
Primary Education Graduates	The fraction of the employed population that have completed primary education; at regional level (NUTS 3). See Appendix B for calculations.	Hellenic Statistical Authority, Educations Statistics & authors' calculations
Secondary Education Graduates	The fraction of the employed population that have completed secondary education; at regional level (NUTS 3). See Appendix B for calculations.	Hellenic Statistical Authority, Educations Statistics & authors' calculations
Tertiary Education Graduates	The fraction of the employed population that have completed tertiary education; at regional level (NUTS 3). See Appendix B for calculations.	Hellenic Ministry of Education, & authors' calculations
Employment Density	Employment divided by square kilometres; at regional level (NUTS 3).	Hellenic Statistical Authority, Social Accounts Section & authors calculations
Labor Force Density	Labour force divided by square kilometres; at regional level (NUTS 3).	Hellenic Statistical Authority, Social Accounts Section & authors calculations
Population Concentration	$\left(\sum_{i=1}^{N} \lvert p_i - a_i \rvert \big/ 2\right) * 100$ where p_i is the population share of region *i* or prefecture, α_i is the area of region or prefecture *i* as a percentage of the country area, N stands for the number of regions and $\lvert \ \rvert$ indicates absolute value. The index lies between 0 (no concentration) and 100 (maximum concentration) in all regions.	Hellenic Statistical Authority, Social Accounts Section & authors calculations

TABLE 2: DESCRIPTIVE STATISTICS

	OBS.	MEAN	ST. DEV.	MIN	MAX
GDP per worker	1173	27,347.11	7,890.87	14,953.00	85,456.00
Public Capital per worker	1173	72,192.15	186,829.80	5,606.00	1,646,179.00
Primary Education Graduates	1173	0.44	0.09	0.14	0.62
Secondary Education Graduates	1173	0.28	0.10	0.07	0.53
Tertiary Education Graduates	1173	0.11	0.04	0.04	0.24
Employment Density	1173	25.47	48.58	3.00	432.30
Labor Force Density	1173	27.93	52.07	3.70	472.70
Population Concentration	1173	40.55	0.83	39.36	41.82

CONSTRUCTION OF EDUCATION DATASET

We compile a dataset which disaggregates the employed population into the three main education levels, namely primary, secondary and tertiary education graduates, for the 51 NUTS 3 Greek regions and the period 1981-2003. We define the primary school graduates as a percentage of employment for each year *t*, as follows:

$$PE_t = (\textit{Employed Primary Education Graduates}_{t-1} + \textit{New Employed Primary Education Graduates}_t - \Delta(\textit{Retirees with Primary Education})_t - \textit{New Employed Secondary Education Graduates}_t)/\textit{Employment}_t \qquad (A1)$$

So, primary school graduates employed in each year equal previous year's employed graduates plus new graduates entering employment minus graduates exiting employment the current year. The graduates exiting employment are composed of those who retire and the employed who obtain the qualification just above primary school (i.e. secondary school) the present year. In order to obtain the stock of primary school graduates who work every period, we first take the employed primary school graduates from the relevant Census year available (1981, 1991, 2001). Second, we estimate this portion for the remaining years of 1981-2003 by interpolation and extrapolation using the above calculations.[18]

[18] For interpolation we use the *csipolate* Stata command, which fills the gaps of missing values by averaging non-missing values using cubic spline interpolation (Hamming, 1973; Press et al., 2007). For extrapolation, we use the *ipolate* Stata command, which extends a series using linear methods (Stata Reference Manual, Release 11, volume D, 2009).

Regarding *New Employed Primary School Graduates*, since we do not have yearly data on them, we calculate them for *t* as the difference in primary school enrollments between years *t-6* and *t-7,* given that the duration of primary school in Greece is 6 years, minus the percentage of students who did not finish primary school this period. We calculate the latter percentage by subtracting the number of students who finished primary school from those who enrolled in primary school but did not finish it in *t*. We precisely calculate this percentage for 1981, 1991 and 2003 and interpolate it for the intermediate years. Afterwards, we adjust this number by multiplying it by the fraction of the working population with primary education. We calculate this fraction for all Census years (1981, 1991, 2001) and interpolate or extrapolate it for the remaining years of 1981-2003. Please note that school enrolment data for primary and secondary education are available for all regions and years from the Education Statistics, Hellenic Statistical Authority.

To calculate new retirees, we subtract employed population over 65 in *t-1* from the employed population over 65 in *t* with primary education. We calculate these numbers for each region by multiplying Greek population with primary education over 65 by the portion of the Greek population with primary education over 65, corresponding to the respective region. We further subtract from this difference the employed individuals over 65 with primary education who stop working between *t* and *t-1*, taking the difference of employed over 65 with primary education between *t* and *t-1*. We classify population over 65 and employed over 65 into different education levels using Census data for 1981, 1991 and 2001 and interpolate or extrapolate for the remaining years of 1981-2003.

The $New\ Secondary\ Education\ Graduates_t$ are calculated taking the difference in secondary school enrollments between periods *t-6* and *t-7* minus the percentage of students who did not finish lower secondary school in *t*, since the duration of secondary school in Greece is 6 years. We calculate the latter percentage by subtracting the number of students who finished the secondary school grade in which they enrolled from the enrolled students who did not finish it in *t*. We calculate this for 1981, 1991 and 2003 and interpolate it for the intermediate years of 1981-2003. We also use total employment data for each Census year and interpolate or extrapolate it for the remaining years of the period 1981-2003 in order to get $Employment_t$. After these intermediate steps, we calculate *Employed Primary School Graduates* as a percentage of total employment for each year using *(A1)*.

We calculate the secondary and tertiary education graduates as percentages of employment in a similar way. Thus:

SE_t = *(Employed Secondary Education Graduates*$_{t-1}$ + *New Employed Secondary Education Graduates*$_t$ - *Δ(Retirees with Secondary Education)*$_t$ - *New Employed Tertiary Education Graduates*$_t$*)/Employment*$_t$. *(A2)*

We define *Tertiary Education Graduates* as a percentage of employment as follows:

TE_t = *(Employed Tertiary Education Graduates*$_{t-1}$ + *New Employed Tertiary Education Graduates*$_t$ - *Δ(Retirees with Tertiary Education)*$_t$*/Employment*$_t$ *(A3)*

There is one difference between the calculations of *Employed Tertiary Education Graduates* with respect to the calculation of *Employed Primary Education Graduates and Employed Secondary Education Graduates* as shares of total employment. In the numerator of the definition of the first variable, we do not subtract employed people with higher than tertiary education qualification in *t*, as we do for the other two variables, since tertiary education is the highest education level we consider. The *New Tertiary Education Graduates*$_t$ are calculated taking the difference in tertiary school enrollments between periods *t-6* and *t-7* since the duration of tertiary education in Greece is between 4-6 years. Please note that university enrolment data are available for all regions and years from the Hellenic Ministry of Education.